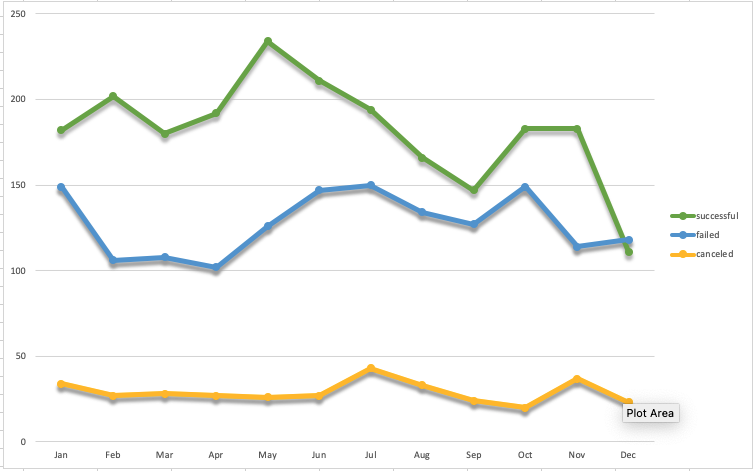
Excel Homework: Kickstart My Chart

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?

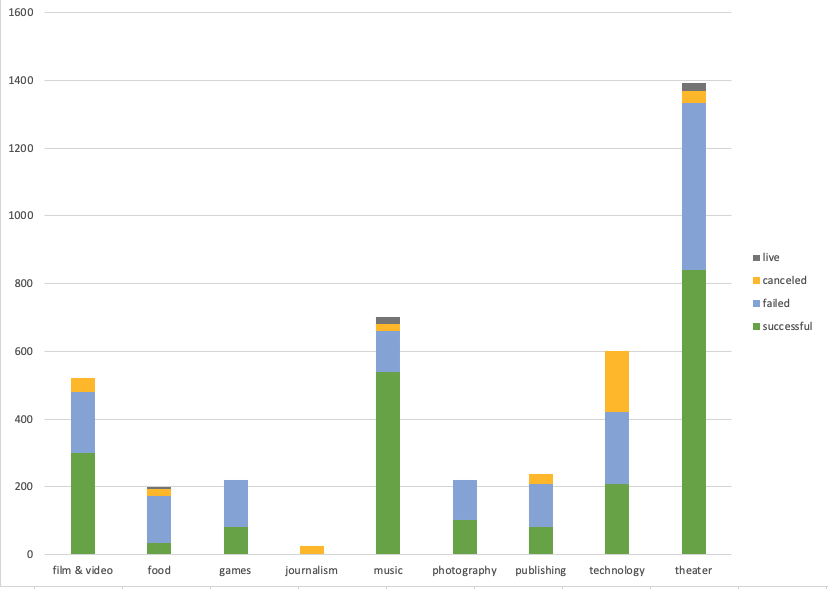
We built three different charts, but I think we can see the same conclusion that campaigns are dropping the successful rate and failed campaigns rate are increasing.

Why this thing happened? We can see the time series graph.

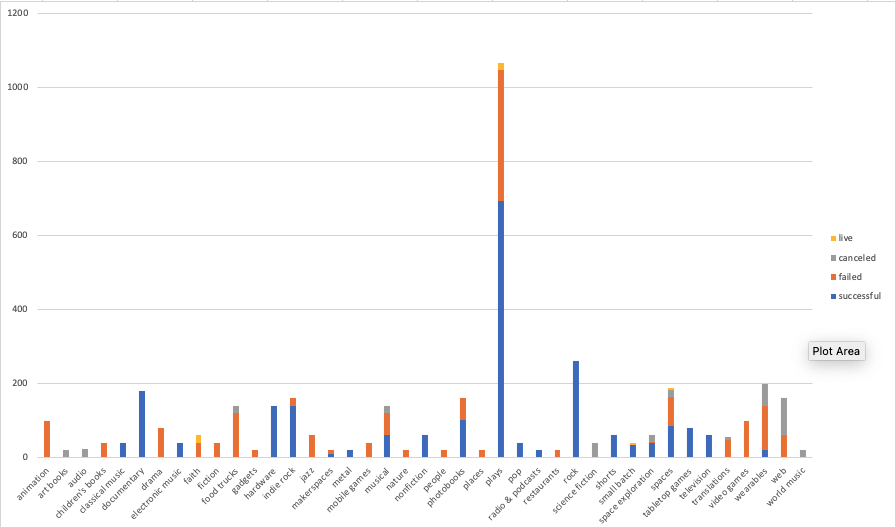


This graph show that successful rate is decreasing through a year. We can assume that the customers were curious at the beginning but as time flows losing.

How can we change this situation, we can see other graph.



First, campaign categories show us the theater category has the highest successful numbers but high failed at the same time. If we want to drop down the risk of failed rate, we should choose the music category which has high successful rate rather than others.



Second, if we go more detail about the campaign categories in this chart, we can know same as previous. Plays category also has high successful number and failed at the same time. But documentary, classical/electronic music, pop or rock only have successful events.

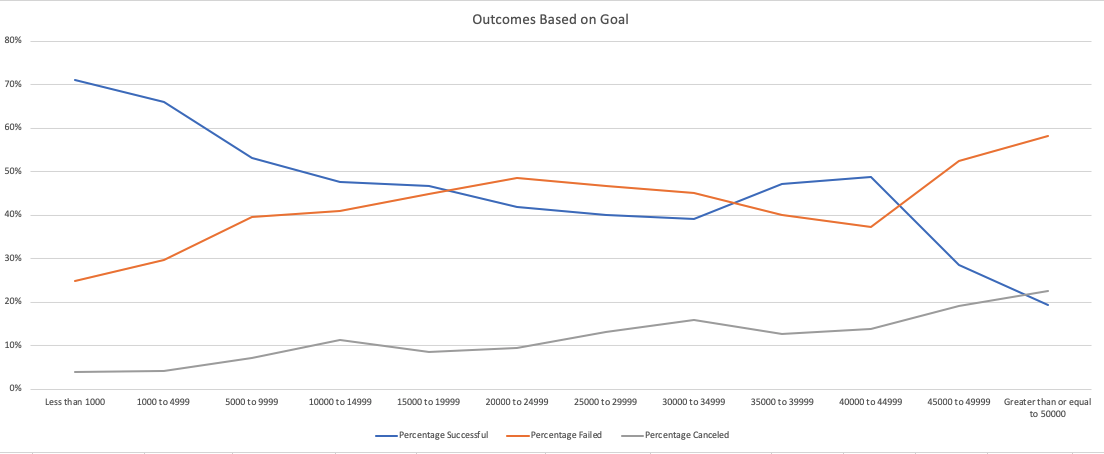
This we can focus on this sort of campaign and increasing the successful rate next season.

1. What are some limitations of this dataset?

This dataset doesn’t have more detail information like cost of campaign, the feedback of campaign or funder information. The more data we have, the more analysis we can do. What type of campaign funders are like? How much we need to cost to hold a campaign? How was the customer feel about the campaign, etc. I think we need more specific information that can help our analysis.

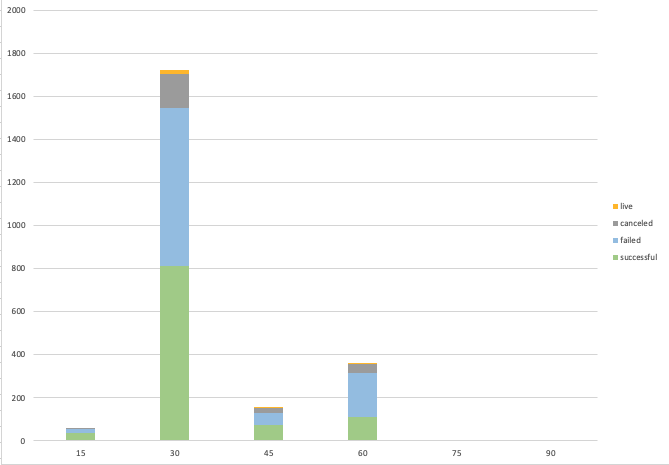
1. What are some other possible tables and/or graphs that we could create?

Like the Bonus part, we built the percentage state on number of goal range. And we can know our campaign only get high percentage of success when we hold a campaign which benefit less than $1000 and high failed rate on greater than $5000.



I think this is the first problem we need to face. Or we will keep losing money and funder interest.

We also can use “Date Created Conversion” and “Date Ended Conversion” columns to build the line chart. Show the correlation between state and long of campaign.



As you can see this chart, the most common events are a month campaign, but as day increase, the number of failed would go up. Lots of 60 days campaign are failed. We need to avoid long term campaign and combine the conclusion we made before.

Focus on hold music category campaigns less than 30 days in order to increasing successful rate before next season.

Table

Description automatically generated

Chart, bar chart

Description automatically generated

1.Use your data to determine whether the mean or the median summarizes the data more meaningfully.

As the picture, for the successful campaigns, the median summarizes the data more meaningfully. On the other hand, for the failed campaigns, the mean and median both are not meaningful to the data set.

2.Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

I think in this dataset the state column is the most meaningful data we can use, because it can tell us if campaigns successful or not. But how do you decide successful campaigns or not? The “state” column we use now, is based on pledged numbers more than goal, but it’s only short term goal, if the campaigns can bring long term benefit, it might have other definition for state column.

Therefore, I think the most important data we might can have in the future is customer feedback, and how to quantize data is the challenge for the data analyst.

*Chiaching,Hsieh (Joyce) 2021/06/14*